

## FREEDOM, DESIRE, AND NECESSITY

### AUTONOMOUS ACTIVITY AS ACTIVITY FOR ITS OWN SAKE

*Pascal Brixel*

IN THE *Nicomachean Ethics*, Aristotle takes seriously only three candidates for the good human life: the life of pleasure, the life of politics, and the life of contemplation. These are the serious contenders because each of them has some claim to be not merely instrumentally good but desirable for its own sake:

If ... there is some end of the things we do, which we desire for its own sake (everything else being desired for the sake of this), and if we do not choose everything for the sake of something else (for at that rate the process would go on to infinity, so that our desire would be empty and vain), clearly this must be the good and the chief good.<sup>1</sup>

Hannah Arendt argues that Aristotle is here guided not just by a view of the good but by a view of freedom.<sup>2</sup> For he rejects other ways of life, such as “the life of money-making,” on the basis that they are “undertaken under compulsion,” wealth in particular being “merely useful and for the sake of something else.”<sup>3</sup> Elsewhere, he says that wealth acquisition “may be studied by a free man, but will only be practiced from necessity,” and that the occupations it encompasses are “the most servile” arts.<sup>4</sup> And he is similarly dismissive of the life of labor aiming at the immediate sustenance and reproduction of life itself—cooking, eating, procreating, cleaning, etc. In the *Ethics* he simply ignores it, and in the *Politics* it appears as the life proper to slaves and animals: “both with their bodies administer to the needs of life.”<sup>5</sup>

The compulsion that Aristotle takes such activities to involve is not involuntariness in some inner sense, such as unawareness of what one is doing, as in

1 Aristotle, *Nicomachean Ethics*, 1094a.

2 Arendt, *The Human Condition*.

3 Aristotle, *Nicomachean Ethics*, 1096a.

4 Aristotle, *Politics*, 1258b.

5 Aristotle, *Politics*, 1254b.

sleepwalking, or capitulation to overpowering psychological forces, as perhaps in some addictions or phobias. Nor, notwithstanding the reference to slavery, is it fundamentally a matter of subjection to interpersonal coercion. It is rather the compulsion that consists in having to do something merely as a means—out of mere necessity for an end external to the activity. A truly free activity, on this view, is one chosen for its own sake and not merely for the sake of further ends.

A version of Aristotle's idea appears again in the work of Marx. In volume 3 of *Capital*, Marx famously identifies "the true realm of human freedom" with "the development of human powers as an end in itself," which he opposes to "the realm of necessity," or of "labor determined by necessity and external expediency."<sup>6</sup> Marx takes a more expansive view of the human good than Aristotle: the highest good consists no longer in pure contemplation but in the development and exercise of the totality of a human being's powers. But Marx shares Aristotle's sense that a truly free activity cannot proceed from mere instrumental necessity. Free activity must be, at a minimum, activity for its own sake.

Despite its pedigree, this Aristotelian-Marxian idea about freedom has not found its way into mainstream discussions of freedom and autonomy in contemporary analytic philosophy, even as an object of criticism. Perhaps this is because neither Aristotle nor Marx ever developed or defended the idea in a systematic, comprehensive, and fully explicit way. Be that as it may, I will argue in this paper that it constitutes a serious contribution to thought about autonomy and promises to solve important and abiding problems in the contemporary literature on that topic.

Autonomy—freedom, in one traditional sense of the word—is self-determination.<sup>7</sup> To act autonomously is to be determined by one's own will rather than by alien compulsions. Since such compulsions can take multiple forms, autonomy has multiple dimensions. In its internal dimension, autonomy is opposed to the psychological unfreedom that stems from phobias and addictions, maladaptive preferences, brainwashing, and the like. In its external

6 Marx, *Capital*, vol. 3, 958–59.

7 Throughout the paper, when I speak without further specification of "autonomy," I mean "local" rather than "global" autonomy. That is, I am interested in what it is for a particular activity to be autonomous, rather than what it is for a person to be autonomous over the course of their whole life. Put another way, I am interested in what it is for someone to *do something autonomously*, rather than what it is for someone to be autonomous *simpliciter*. Note also that I will have nothing to say about various important adjacent issues, such as the role of autonomy in a specifically political ideal, or whether there is a right to autonomy, or under what conditions we can hold people morally or legally responsible for their doings. Approaches to autonomy that are motivated primarily by concerns about rights and responsibility are thus outside the scope of this paper. Perhaps it will turn out that little progress can be made if we put these concerns aside. But let us see.

dimension, it is opposed to the unfreedom associated paradigmatically with coercion by other people and duress by unfavorable circumstances.

My concern here is with autonomy in its external dimension.<sup>8</sup> Our ordinary thought about this dimension is structured by two conceptual oppositions. First, as mentioned, autonomy is opposed to external unfreedom; that is, both coercion and duress compromise autonomy. Second, whereas to act autonomously is in some sense to do what one wants or desires, to do something under coercion or duress is in some sense to be forced or necessitated to do it. An adequate theory of external autonomy, I argue, ought to make sense of both of these platitudes. It ought to tell us why coercion and duress compromise autonomy, and it ought to identify the relevant senses of desire and necessity that structure our ordinary thought about these matters.

In the first half of the paper, I defend these theoretical desiderata and show that standard contemporary accounts of autonomy struggle to deliver the requisite explanations. I then defend an alternative necessary condition of autonomous activity along Aristotelian-Marxian lines. One does something autonomously, I argue, only if one does it for its own sake and not *merely* for the sake of a further end. This idea turns out to steer an attractive middle path between two dominant families of contemporary theories of autonomy. In this way, I argue, it puts us in a position to explain how coercion and duress compromise autonomy, and to articulate the sense in which to act autonomously is to act under the guise of desire, while to act unfreely is to act under the guise of necessity.

## 1. AUTONOMY AND EXTERNAL UNFREEDOM

I begin in this section by identifying two pieces of common sense that a complete theory of autonomy should try to respect and indeed explain. First, coercion and duress can compromise autonomy. Second, while autonomous activity is, at a minimum, a matter of doing what one really or truly *wants* to do, victims of coercion or duress are *forced* to act as they do—where the relevant kinds of desire and necessitation are conceptually opposed to each other.

### 1.1. Coercion and Duress

If autonomy is self-determination, it is hard to deny that coercion by another person compromises autonomy. When I speak of coercion, I have in mind not the direct use of force against someone's body but rather what is sometimes

8 I will thus put aside, for instance, the question of how to distinguish those kinds of internal compulsion that make us unfree from those that do not. An example of the latter might be the sort of inner compulsion from which Martin Luther is said to have acted in declaring, "Here I stand; I can do no other."

called “volitional” coercion, which operates by means of threats. Think of the slave who does her master’s bidding to avoid a beating, or the mugging victim who hands over her money to avoid being shot. Action of this kind is a paradigm case of heteronomy, determination not by one’s own will but by an alien compulsion, in the form of another’s will.

Though coercion is certainly the most obvious and egregious type of external unfreedom, it is very natural to think that autonomy can also be compromised by duress. When I speak of duress, I have in mind compulsion by unfavorable circumstances that force an agent to choose the lesser of two evils, where the relevant circumstances do not result from another’s deliberate machinations. Think of a sailor who throws goods overboard in a storm to save herself and her ship, or someone who abandons her house to escape a fire. Like (volitionally) coerced actions, these actions are intentional. But like coerced actions, they fall short of unqualified self-determination, though the alien compulsion in this case originates not in another will but in the agent’s exigent circumstances.

No doubt there are deep and important differences between interpersonal coercion and impersonal duress, with respect to both their nature and their moral implications. No doubt our interest in (being free of) each is motivated to some extent by concerns distinctive to each. But equally, there certainly appears to be a generic unity underlying these specific differences. As I have observed, it is perfectly natural to speak of both coercion and duress as sources of unfreedom, inasmuch as both seem to compromise self-determination. If possible, a complete account of autonomy ought to vindicate and explain this appearance.

### *1.2. Desire and Necessity*

In pursuing this explanatory task, we should attend to another piece of common sense. Our ordinary thought about coercion and duress alike is structured by a conceptual opposition between a sort of desire and a sort of necessity. Whereas autonomous activity is an expression of the agent’s desire, its opposite—heteronomy, unfreedom—is a reflection of necessitation alien to such desire.

Thus, it is natural to think that victims of coercion and duress act non-autonomously precisely because, in some salient sense, they do not act as they really want to, but are instead forced, compelled, or necessitated to act as they do. Indeed, the fact that we readily affirm this sort of thought about both the interpersonal and the impersonal cases lends further, more concrete support to the suggestion that coercion and duress have something in common, and that both are opposed to autonomy.

The problem, though, is to explain the relevant kind of desire, as well as the distinctive kind of necessity that is opposed to it. The solution is not obvious. After all, there is a straightforward sense in which people acting under coercion

or duress do just what they want to do. Given their nonideal circumstances, the mugging victim handing over her money, the slave doing her master's bidding, the sailor throwing goods overboard, and the person fleeing her house all act intentionally, and presumably they all prefer this action to any of the available alternatives. Now, one wants to say: these agents do not *really* or *truly* want to do these things. But the point is to explain what this means. What we can say at this stage is that the kind of desire that makes an activity autonomous must amount to something more robust than mere intention.

On the flip side, there is an equally straightforward sense in which people acting under coercion or duress are not literally forced to act as they do. The slave could opt to be beaten, and the mugging victim, the sailor, and the person fleeing her house could all choose to take their respective goods to the grave rather than part with them. So the kind of necessitation that is the mark of external unfreedom must amount to something weaker than the absolute impossibility of doing otherwise.

Coercion and duress compromise autonomy; to do something autonomously is in some sense to do what one wants to do, but victims of coercion and duress do what they are forced to do, where the relevant kind of necessitation is opposed to the relevant kind of desire. These platitudes function as theoretical constraints. We have reason to be dissatisfied with a theory that does not vindicate and explain them. But as we will see, they also function as useful footholds. We can try to make progress by starting with these fairly abstract ideas and giving them greater determinacy.

## 2. SUBJECTIVISM AND OBJECTIVISM

In this section, I argue that standard theories of autonomy struggle to explain the platitudes I just described. Since my subject matter is autonomous activity, I focus in the first instance on theories of "local" rather than "global" autonomy: that is, theories of what it is to do a particular thing autonomously, rather than what it is to be an autonomous person over the course of one's life as a whole. These theories tend to be deeply subjectivist, making the problem with coercion and duress fundamentally a function of the agent's reflective attitudes. At the end of the section, I also consider an alternative, more objectivist approach that is more usually associated with global autonomy. The two types of approach, I argue, struggle for opposite reasons. The subjectivist approach is so fundamentally dependent on the agent's subjective attitudes that it leaves us unable to do justice to the role that the concept of autonomy can and should play in the justification and criticism of those attitudes. The objectivist approach, on the other hand, is so divorced from the agent's subjective attitudes that it leaves us

unable to capture the sense in which the unfree agent does not what she wants but what she is forced to do.

### 2.1. Subjectivism

What does the autonomy of an activity require beyond mere intention? If autonomy is self-determination, it seems that the desire from which the agent acts must be internal rather than alien to her: it must be *her own*, in some especially robust sense. To put the point another way, the desire must express her practical identity, or who she is as an agent. Only then does the realization of this desire count as *self-determination*. But what is it for a desire to be robustly “one’s own,” or to express one’s “practical identity”?

The most influential approach to this question—pioneered by John Plamenatz and later developed by Gerald Dworkin, Harry Frankfurt, and many others—centers on our capacity for self-reflection: our capacity to turn our attention toward our own motives and to form higher-order attitudes about them.<sup>9</sup> We are not always moved by our desires automatically, as it were, but have the power to step back from at least some of them and appraise them critically. On the basis of such appraisal, whether in the light of reason or in the light of other elements of our psychology, we can endorse or disavow our first-order desires, approve of being moved by them, or resent being moved by them. But to endorse or approve of one’s first-order desire is in a sense to identify with it, to recognize it as truly one’s own. To disavow or resent it, on the other hand, is to be alienated from it, to look upon it as a force external to one’s true will.

On this account, then, it is the agent’s higher-order attitudes that determine whether a given desire expresses her practical identity or constitutes something more like an alien force acting upon her. It is her higher-order attitudes that make the difference between autonomy and heteronomy. To quote Plamenatz, whose version of the view is especially clear and simple: “freedom must be defined, in its primary . . . meaning, as action from a motive from which a man desires to act, or, at least, does not desire not to act, and the lack of freedom as action from a motive from which he does not desire to act.”<sup>10</sup> The suggestion is thus that an agent who gives in to an unwanted addiction, for example, acts non-autonomously because, although she may do what she desires, the relevant desire is not her own in the fullest sense, because she does not identify with it; on the

9 Plamenatz, *Consent, Freedom and Political Obligation*, ch. 5; Dworkin, “Acting Freely”; and Frankfurt, “Freedom of the Will and the Concept of a Person,” “Identification and Wholeheartedness,” and “The Faintest Passion.” Some in this literature speak of “freedom,” some of “autonomy,” and some simply of “identification.” But beneath the superficial diversity of expression lies an obvious unity of subject matter and content.

10 Plamenatz, *Consent, Freedom and Political Obligation*, 122.

contrary, she wishes that she did not have it, or at least that it did not move her to action on this occasion. She resents being moved to action in this particular way.

There are many sophisticated and idiosyncratic variations on this approach to which I cannot do justice here.<sup>11</sup> Broadly speaking, the variations tend to modify the basic view along three dimensions. First, theories in this tradition vary in terms of what specific types of reflective attitude play the decisive role in identifying an agent with, or alienating her from, her own motives. Some speak simply of higher-order desires; others employ richer notions such as “resentment”; others still use more technical terms such as “commitment” and “policy” (for example, about what to treat as a reason for action). Second, theories vary, too, in terms of the objects of these reflective attitudes. In some versions, the relevant reflective attitudes are attitudes toward one’s own desires. In others, the reflective attitudes instead (or additionally) concern the historical process by which one’s desire came to be. In others still, the (or another) relevant object of reflective attitudes is the choice situation that gave rise to the desire. Finally, many theories stipulate some kind of cognitive condition, to the effect that what really matters is the reflective attitude that the agent would have toward the relevant aspect of her motivational structure *if*, for instance, she were minimally rational and fully informed. Typically, however, these conditions of rationality and the like are quite minimal; to the extent that they are not, the view begins to depart in a more substantial way from the basic Plamenatzian model.

What unites all of these versions of the approach is a deeply subjectivist orientation, in the sense that they all conceive of autonomy as largely a function of what the agent thinks and feels, or would think and feel, about elements of her motivational structure or the situation that produces this structure. The relevant reflective attitudes ultimately divide the agent’s first-order desires into “authentic” desires that are truly her own, and “inauthentic” desires that are hers only in a qualified way, whether because they are immediately objects of a relevant negative reflective attitude or because they arise from a process or from a situation that is the object of such an attitude (or would be if the agent were aware of it). Fundamentally, what makes an activity fully autonomous is that it is an expression of authentic rather than inauthentic desires in this sense.

How can the subjectivist picture help us to understand what is wrong with coercion and duress? For the sake of illustration, let us work in the first instance

11 See, e.g., Neely, “Freedom and Desire”; Thalberg, “Hierarchical Analyses of Unfree Action”; Christman, “Autonomy and Personal History”; Ekstrom, “A Coherence Theory of Autonomy”; Bratman, “Identification, Decision, and Treating as a Reason”; Velleman, *The Possibility of Practical Reason*, ch. 1; J.S. Taylor, “Autonomy, Duress, and Coercion,” 138; Christman, “Procedural Autonomy and Liberal Legitimacy” and “Introduction”; and Killmister, *Taking the Measure of Autonomy*.



with Plamenatz's straightforward version of the view, according to which autonomous activity requires that one does not act from an "undesired" motive, that is, a motive from which one actively desires not to act. This condition seems to deal in an obvious way with some cases of coercion—namely, those involving what the law calls an "overborne will." An agent may divulge a secret under torture to avoid further suffering, while wishing that she had the inner strength to take the secret to her grave. This agent suffers what Frankfurt calls an "inner defeat."<sup>12</sup> She does not do what she *really* or *truly* wants—she acts "against her own will"—inasmuch as she acts from a motive from which she desires not to act. She therefore acts non-autonomously.

Not all volitional coercion involves an overborne will. The victim of a mugging may not suffer an "inner defeat" like the victim of overbearing torture. She may not wish that she had the strength of will to take her money to the grave. Given the circumstances, she presumably not only wants to hand over the money, but very much wants to have this desire and to act from it. Her first-order desire is not straightforwardly an object of her aversion in the sense that she would prefer not to be moved by it, *given* the circumstances.<sup>13</sup>

Even so, however, we can say that she acts from an "undesired" motive in the sense that there is some description of her motivational structure under which she has a negative attitude toward this structure. For instance, it is often suggested that people resent acting not in order to improve their condition but merely to keep it from becoming worse, or that they resent being motivated by "duress," or by "menacing" or "coercive" situations.<sup>14</sup> In any case, what makes action under coercion or duress unfree is that the agent has some negative attitude toward some distinctive aspect of the way in which she is being moved to action. It is because of this negative attitude that an activity that flows from the relevant kind of motivational structure is not a true expression of its agent's will but alienated, importantly foreign to her will.

## 2.2. *The Wrong Order of Explanation*

The notion that victims of coercion and duress resent or are otherwise psychologically alienated from the way they are motivated is supposed to capture the sense in which the relevant desires are not really or truly their own—which in turn explains why these agents act unfreely. But reflective attitudes, such as the

12 Frankfurt, "Coercion and Moral Responsibility," 81.

13 This is the point of Irving Thalberg's objection in "Hierarchical Analyses of Unfree Action."

14 Dworkin, "Acting Freely," 377; Frankfurt, "Coercion and Moral Responsibility," 82–83; J. S. Taylor, "Autonomy, Duress, and Coercion," 138; Thalberg, "Hierarchical Analyses of Unfree Action," 217–18; Christman, "Introduction," 8–9. Compare Killmister, *Taking the Measure of Autonomy*, 109–10.



desire not to act from certain types of motive, cannot play the explanatory role demanded of them in this account, for, properly understood, such attitudes are *responses* to unfreedom rather than *grounds* of it.

Consider, for a start, what we should say if a victim of coercion or duress lacks the relevant kind of reflective aversion to her own motive. Imagine a Stoic slave as described in the discourses of Epictetus. This slave feels no aversion at being threatened and ordered around. When ordered to hold her master's chamber pot, she obeys with equanimity, reasoning simply that "it is a more valuable thing to get a dinner than not; and a greater disgrace to be given a thrashing than not to be."<sup>15</sup> Her lack of reflective aversion is not the product of an impaired capacity for reflection. It is not even the product of a socially inculcated sense of her own inferiority. Rather, it springs from her reflective and authentic moral conviction that she should not rail against that which she cannot control. Does this person act autonomously?

As several philosophers have argued, subjectivist theorists of autonomy must say yes, since from the point of view of the slave's reflective attitudes, everything is in order.<sup>16</sup> But this conclusion is difficult to take seriously. The Stoic slave may indeed feel or believe that she is free. Feeling free, however, or believing oneself to be free, even when this feeling or belief is formed in a minimally rational and fully informed way, is not the same as being free. When the slave holds the chamber pot for her master—something she has no independent desire to do, and which she does only to avoid being beaten—she acts not autonomously but under the compulsion of an alien force. This is a paradigm case of heteronomy.<sup>17</sup>

Superficially, this looks like a problem concerning the extensional adequacy of the theory, and that is how it has tended to feature in the literature.<sup>18</sup> The real problem, however, is more fundamental; it concerns the *structure* of the subjectivist explanation. What the case of the contented slave highlights, I think, is that when we feel aversion at being coerced, for instance, this is not a brute psychological fact about us. For the contented slave does not differ from a slave who resents her condition simply in being constituted psychologically *differently*. Rather, there is something that the latter gets right and the former

15 Epictetus, *The Discourses of Epictetus*, bk. I, ch. 2.

16 Zimmerman, "Making Do"; Kristinsson, "The Limits of Neutrality"; J. S. Taylor, "Autonomy, Duress, and Coercion," sec. 4. As Kristinsson points out, some subjectivists seem to embrace this conclusion quite explicitly; see, e.g., Dworkin, "The Concept of Autonomy," 23; and Christman, "Liberalism and Individual Positive Freedom," 359.

17 For a more thorough discussion of counterexamples in the same vein, see Oshana, "Personal Autonomy and Society" and *Personal Autonomy in Society*, ch. 3.

18 For some responses to the charge of extensional inadequacy, see Killmister, *Taking the Measure of Autonomy*, esp. ch. 6.

gets wrong. That is, the former fails to feel a kind of aversion that her situation warrants, indeed demands. Aversion to being coerced—aversion at being motivated in this kind of way—is thus not a brute psychological response to a stimulus but a *warranted response* to this kind of treatment.

But what is it that warrants this negative reflective attitude? In other words, what is objectively wrong with being subjected to coercion or duress, such that this condition normally and properly gives rise to reflective aversion? Most obviously, what is wrong with it is precisely that to act in this way is to act unfreely. Such action is not self-determined, and this is why we can be expected to do it only with aversion. What the Stoic slave fails to recognize, then, is that her condition warrants or demands aversion because it is a condition of unfreedom. Her equanimity is a mark of insufficient concern for her own autonomy.

From a subjectivist point of view, however, we cannot make sense of these natural and attractive thoughts. For from that point of view, as Dworkin says quite explicitly, “we do not find it painful to act because we are compelled; we consider ourselves compelled because we find it painful to act for these reasons.”<sup>19</sup> We can now see that this gets things the wrong way around. The various modifications of the basic subjectivist approach do not help with this problem. For instance, it makes no difference here whether what a victim of coercion or duress reflectively resents is the motive from which she acts or the process or situation that gives rise to this motive. In any of these cases, we are owed an explanation of why she happens to resent just *these* kinds of motive, process, or situation and not others, and this explanation should make the resentment intelligible as a *warranted response* to the relevant kind of motive, process, or situation. But the prospects for a successful explanation from a subjectivist point of view are dim, given the unavailability of the most obvious such explanation—namely that the relevant kind of motive, process, or situation makes us unfree.

The Plamenatz-Dworkin-Frankfurt approach to understanding the unfreedom of action under coercion and duress, then, is not promising on reflection. Because of the depth of its subjectivist dependence on the agent’s psychology, this approach seems to leave us unable to account adequately for the crucial role that the concept of autonomy itself can and should play in the justification and criticism of our reflective attitudes. The problem is not simply that we sometimes fail to feel resentment when we act under coercion or duress. It is that when we do feel such resentment, it is normally, properly, and fundamentally because coercion and duress are making us act unfreely—because they involve forms of motivation that compromise our autonomy. But then it cannot

19 Dworkin, “Acting Freely,” 378–79. Compare Epictetus: “Nobody, then, who lives in fear, pain, or distress is free; but whoever is delivered from pains, fears, and distresses, by the same means is delivered likewise from slavery” (*The Discourses of Epictetus*, bk. II, ch. 1).

be the case, conversely, that they compromise our autonomy fundamentally because we resent them. Resentment is downstream of unfreedom.<sup>20</sup>

### 2.3. *The Wrong Kind of Necessity*

In our ordinary thought, the sense in which the unfree agent does not do what she really or truly wants is opposed to a correlative sense in which she is instead forced to do it. A secondary shortcoming of the subjectivist approach is that it suggests no plausible account of the relevant kind of necessity.

To act from a motive that is an object of one's aversion is not in itself to act under the guise of compulsion or necessity. The mugging victim may choose to hand over her money merely in order to avoid the greater evil of being shot but, strictly speaking, she could have chosen instead to take the money to her grave. Whether or not she happens to feel aversion at the motive from which she in fact acts simply seems to be neither here nor there, as far as the necessity or non-necessity of her action is concerned. Of course, since people cannot be expected voluntarily to act from motives to which they are averse, it stands to reason that she must indeed be forced to act as she does. But we already knew this. The problem was, and still is, to explain what it means and why it is true.

Perhaps a further proposal from Frankfurt could help. In "Coercion and Moral Responsibility," he suggests that the desires to which coercion and duress give rise in their victims are "irresistible." The idea is that the victim of a mugging, for instance, will typically not only acquire a desire to acquiesce to the mugger's demands so as to save her life, but will find it psychologically impossible not to act on this desire, even if she wants to resist it.<sup>21</sup> "The victim's desire or motive to avoid the penalty with which [she] is threatened," Frankfurt says, "is—or is taken by [her] to be—so powerful that [she] cannot prevent it from leading [her] to submit to the threat."<sup>22</sup> The kind of necessity that characterizes external unfreedom, on this view, is the inexorability with which the desire leads to its own expression in action.

Though Frankfurt at least takes the problem of necessity seriously, his proposal looks like a step in the wrong direction. The suggestion that victims of

20 The same problem arises for idealized versions of the subjectivist picture. Thus, Robert Nozick explains the unfreedom of coerced action by appeal to the aversion that the "Rational Man" feels at being threatened ("Coercion," 460). But of the Rational Man, too, we should be able to say that he feels aversion at being threatened *because* this makes his action unfree.

21 Frankfurt, "Coercion and Moral Responsibility," 80–81.

22 Frankfurt, "Coercion and Moral Responsibility," 77. For the sake of simplicity, I am largely ignoring Frankfurt's caveat that the victim might merely *think* that she is unable to resist the desire. I do not think this makes a difference to my arguments.

coercion generally act from desires that are literally irresistible wrongly assimilates all cases of coercion to those special cases that involve an overborne will.<sup>23</sup> A mugging victim might happen to be so frightened by the prospect of imminent death that she is unable to resist the desire to acquiesce. But another victim, perhaps endowed with an unusually proud and rebellious temperament, might find herself pulled more or less equally toward acquiescing on the one hand and resisting her coercer on the other. She might be perfectly *capable* of choosing the option of dying rather than acquiescing. If this victim ends up acquiescing, however, her action, too, is forced and less than fully autonomous.

Relatedly, the appeal to the irresistibility of desire seems to misidentify the source of necessitation in typical cases of external unfreedom. On Frankfurt's proposal, a coerced action is necessitated by the irresistible desire that motivates it. When we speak of coerced actions as forced, however, what we have in mind is that they are necessitated not by the agent's own overpowering urges but by something altogether external to the agent, such as her coercer or her environment. We have yet to make sense of this thought. These first two mistakes—that coercion in general involves an overborne will, and that the source of necessitation is internal to the agent herself—appear to be symptoms of a single error: the attempt to assimilate the external dimension of autonomy too directly to its internal dimension.<sup>24</sup>

Finally, the structure of Frankfurt's account should give us pause. On his view, coercion and duress compromise autonomy because of the confluence of two conditions: the agent's aversion to the desire from which she acts, and the psychological irresistibility of this desire. The former is supposed to explain the sense in which she does not do what she really wants, and the latter the sense in which she is forced to act as she does. But these two conditions are independently intelligible, and their confluence is fundamentally accidental.

By contrast, in our ordinary thought about these matters, the relevant kind of desire and the relevant kind of necessity are internally related, as the poles of a conceptual opposition. When we say that the victim of coercion or duress acts unfreely because she does not do what she really wants *and* because she is forced to act as she does, we are offering not two separate explanations, or even two separate parts of an explanation, but one and the same explanation in two different ways. To make sense of this thought, we will need a different sort of account, one on which there is an internal relation between doing what

23 For a version of this objection, see, e.g., Zimmerman, "Making Do," 38.

24 This is related to James Stacey Taylor's charge that Frankfurt's theory of autonomy is "metaphysical" rather than "political." (J. S. Taylor, *Practical Autonomy and Bioethics*, ch. 3.) See also Oshana, *Personal Autonomy in Society*, *passim*.

one does not really want to do on the one hand and being forced to do it on the other.

#### 2.4. Objectivism

Given that subjectivist theories struggle to make sense of external autonomy and unfreedom, perhaps we should instead try formulating some objective conditions of autonomy that do not depend on the agent's desires or other attitudes. Such conditions have typically been defended in the context of theories of global autonomy—but conceivably they could also help us with our local-autonomy problem.

Joseph Raz argues that a person is autonomous, over the course of her life as a whole, only if she is independent—that is, not subjected to the will of another person—and she has an adequate range of options to choose from.<sup>25</sup> As marks of global autonomy, these conditions seem plausible enough. A person who lives her life under the arbitrary power of another cannot truly be said to be self-governing; she lacks the final authority to choose how to live her life, making such choices only at the whim of the other person. Likewise, we might think that if someone lacks the opportunity to choose how to live her life from among a range of diverse and valuable options, she thereby lacks the opportunity for one important kind of exercise of her agency. In the absence of valuable alternatives to the life she in fact leads, she can say yes to this life but there is a sense in which she cannot say no to it.<sup>26</sup>

I do not doubt that some version of the independence condition is also true as a condition of the autonomy of a particular activity. Someone who acts under coercion is obviously subjected to the will of another person, and this violation of her independence compromises the autonomy of her action. Nevertheless, I

25 Raz, *The Morality of Freedom*, ch. 14, sec. 1. For a similar account, see Oshana, *Personal Autonomy in Society*, ch. 4. For other defenses of something like the independence condition, see, e.g., Pettit, *Republicanism*, ch. 1; Lovett, "Domination"; and List and Valentini, "Freedom as Independence." In a similar vein, Al Mele suggests a non-coercion condition for local autonomy, comparable to the "procedural independence" condition that Dworkin defends in his later work, though Dworkin does not apply it to coercion specifically. (See Mele, *Autonomous Agents*, ch. 10, sec. 4; and Dworkin, "Autonomy and Behavior Control," 25, and *The Theory and Practice of Autonomy*, 18–19. Compare Christman, "Procedural Autonomy and Liberal Legitimacy," sec. 1.) For other defenses of something like Raz's options condition, see, e.g., Cohen, "The Structure of Proletarian Unfreedom"; van Parijs, *Real Freedom for All*, ch. 1; and Goodin et al., *Discretionary Time*, ch. 2. I should reiterate at this point that I am largely putting aside those approaches whose principal aim is to explain why coercion violates rights or undermines responsibility.

26 Hurka, "Why Value Autonomy?"

do not find this condition helpful, at least in the form in which Raz presents it and with respect to the problem I am trying to solve in this paper.

It is worth emphasizing, to begin with, that in the absence of further analysis the condition amounts to little more than a restatement of a theoretical desideratum, or perhaps a placeholder for an explanation. Since, as we saw, the coerced agent does just what she intends to do, pursuing her own ends as best she can in admittedly nonideal circumstances, the difficulty is to understand in what sense she is indeed subjected to the will of another person rather than being directed by her own will. Significant theoretical work, then, remains to be done here. More to the point, though, even a helpfully fleshed-out independence condition will only tell us something about interpersonal unfreedom in particular. That is, perhaps it could ultimately explain what makes interpersonal unfreedom special vis-à-vis other forms of unfreedom—an important problem but one outside the scope of this paper.<sup>27</sup> It will not solve the problem I am trying to solve, which is to explain what interpersonal coercion and impersonal duress have in common such that they are both forms of unfreedom.<sup>28</sup>

On its face, the options condition seems more promising in this respect.<sup>29</sup> Impersonal duress does not violate independence, but is characterized by a lack of valuable alternatives to the chosen course of action. The sailor who must choose between throwing his goods overboard and allowing his ship to sink lacks an adequate range of options to choose from, and this makes his choice less than fully autonomous. And victims of coercion, of course, also frequently lack an adequate range of options. On reflection, however, things are not so clear.

First, I doubt that an adequate range of options, or the availability of valuable alternatives, is necessary for a particular activity to be done autonomously. We generally lack a valuable or even minimally acceptable alternative to eating. In reasonably favorable circumstances, it is true that we can normally choose what and when to eat, but we cannot reasonably choose never to eat. Nevertheless, we do not normally eat unfreely. The lack of a valuable alternative to eating does not in itself compromise the autonomy of this activity. The problem seems to be that the options condition is too disengaged from the agent's desires. For, arguably, the reason we can eat autonomously even in the absence of valuable alternatives is that it is possible for us to really or truly *want* to eat, and that if this desire takes the relevant form (still to be clarified), this may make the activity autonomous.<sup>30</sup>

27 Thanks to an anonymous reviewer for this suggestion.

28 Thanks to Brookes Brown for this suggestion.

29 Thanks to an anonymous reviewer for urging me to consider Raz's options condition.

30 Perhaps we could try a more modest version of the options condition, according to which a range of options counts as adequate so long as one has at least one valuable option. This certainly goes against the spirit of Raz's idea, but if we understand "valuable" as *intrinsically*

Second, the options condition also does little to explain why victims of duress are forced to act as they do. Raz points out that “a choice between survival and death,” for instance, “is no choice from our perspective.”<sup>31</sup> Thus, one who does something merely as a condition of survival has no real choice but to do it, and is effectively forced to do it. Of course we want to say this.<sup>32</sup> But what entitles us to say it? Why is the language of necessitation appropriate? After all, a choice between survival and death actually *is* a choice, and the agent is not literally, absolutely forced to choose survival. Even putting aside the worry about its extensional inadequacy, then, the options condition merely kicks the can down the road.

To be clear, I do not think that the problem I am trying to solve is the problem that Raz is trying to solve—so none of this is intended as a direct criticism of his theory. Nevertheless, it is instructive to consider why his resources prove so unhelpful for my purposes. The root of the problem, it seems to me, is that objectivist conditions of autonomy are not promising when they become, as it were, too objective. The options condition, in particular, does not make reference to the agent’s desires in a way that would help us understand the sense in which a victim of duress does not act as she really or truly wants to. As a result, it becomes very hard to articulate the sense in which her action is genuinely forced and non-autonomous. The fact that such an agent does make a genuine choice—that she does what she intends to do in the pursuit of her own ends—keeps annoyingly rearing its head.

In order to make progress, then, I propose that we return to the dialectic of desire and necessity that structures our ordinary thought about autonomy and external unfreedom. In particular, I propose to start again with the question what kind of desire is necessary for autonomous activity.

---

*valuable*, and we suppose further that the agent chooses the valuable option *on account* of its perceived value, then we arrive at roughly the view I end up defending in section 3.

31 Raz, *The Morality of Freedom*, 376.

32 At least this is common ground between me and Raz. Following James Stacey Taylor, one might argue instead that a victim of impersonal duress, of the sort to which Raz’s options condition might be intended to apply, suffers not a genuine impairment of her autonomy but merely a diminution of the value of her autonomy to her (J. S. Taylor, *Practical Autonomy and Bioethics*, 108). Alternatively, one might argue, following another proposal of Taylor’s, that there is a tripartite distinction at work, so that one can lack autonomy without yet being unfree: perhaps, then, a victim of the relevant kind of duress does lack autonomy but does not act unfreely. As I argued in section 1, however, a theory of autonomy should aim to vindicate both the appearance that impersonal duress genuinely compromises autonomy and the appearance of a bipolar conceptual opposition between the kind of desire associated with autonomy and the kind of alien necessitation associated with unfreedom. While we may be forced to revise these desiderata on reflection, I think we ought to try to satisfy them if possible. Thanks to an anonymous reviewer for pressing me on these points.



## 3. AUTONOMOUS ACTIVITY AS ACTIVITY FOR ITS OWN SAKE

In this section, I introduce and defend a different approach to autonomy and external unfreedom, following the Aristotelian-Marxian line of thought I discussed in the introduction. One does something freely, I argue, only if one does it for its own sake—that is, on account of its perceived intrinsic value—and not merely instrumentally. I motivate this idea on the basis of the same guiding intuitions as the standard subjectivist approach, but I show that it ultimately allows us to steer a promising middle path between subjectivism and objectivism, putting us in a position to explain how action under coercion or duress is non-autonomous.

3.1. *Intrinsic and Extrinsic Desires*

Let me begin with a familiar distinction between different modes of activity. On the one hand, we do some things partly or wholly for their own sake, as final ends. These can include activities, such as some forms of play and contemplation, that do not have any further end at all. But they also typically include some instrumental activities, such as meaningful work. We do not want to live lives of pure leisure; we value intrinsically the pursuit of productive activities that allow us to exercise and develop our skills while making valuable social contributions. On the other hand, there are those things, such as cleaning the toilet, that we do *merely* instrumentally, purely for the sake of further ends distinct from the activities themselves.

These modes of activity correspond to importantly different kinds of desire. A final end is something we do on account of its perceived intrinsic value: value that it has in its own right and that does not reduce to the value of other things, such as its effects.<sup>33</sup> Insofar as we do something for its own sake, then, we have what can be called an “intrinsic” desire to do that thing, for the perceived good that we are after lies in the activity itself. By contrast, instrumental activity is done on account of the perceived value of something else: the further end to which it is a means. Insofar as we do something instrumentally, then, we can be said to have an “extrinsic” desire to do that thing, and insofar as we do something *merely* instrumentally, we have *only* an extrinsic desire to do it.

An extrinsic desire is really a desire only in a qualified way. It is always parasitic on an intrinsic desire for something else, and in a sense it is always this other,

33 Christine Korsgaard argues that to value something as a final end is not necessarily to value it intrinsically (“Two Distinctions in Goodness”). But her argument relies on a gratuitously narrow identification of intrinsic value with unconditional value, value that an object would have in any context whatsoever; see Tannenbaum, “Categorizing Goods.” Thanks to Dan Moerner for drawing my attention to Korsgaard’s and Tannenbaum’s papers.

intrinsic desire that is the agent's "true" desire, the desire whose object is what she is really after. When someone intentionally cleans a toilet, for example, there is an obvious sense in which she "wants" to clean it. But the reason it sounds odd to say this is that her desire to clean the toilet is likely to be purely extrinsic. The real object of value for her—the object of her true and unqualified desire—is likely not the activity of cleaning the toilet but the further ends to which this activity is a mere means, such as living in a hygienic and attractive environment.

With this distinction in hand, we can begin to see more clearly what is at stake in our pursuit of autonomy, in view of familiar considerations of identification and alienation. An autonomous activity, it seems, ought to be a suitably unqualified expression of its agent's will. We said that this means that the agent ought to be able to recognize in the action, and in the desire that motivates it, her own practical identity, or "who she is" as an agent. On the subjectivist picture, this idea was spelled out, unsuccessfully, in terms of a distinction between authentic and inauthentic desires. But we can appeal more fruitfully to our new distinction.

The commitments that are definitive of an agent's practical identity without qualification—the commitments with which she identifies in the first instance—are surely her intrinsic desires, the ends that she pursues for their own sake and not merely for the sake of other things. We care about autonomy because we wish to be the authors of our own lives—but it is our final ends that really determine what we are after in life, what we want our lives to look like. It is when our activity is an immediate realization of our final ends, then, that it expresses our practical identity in the world without qualification. Of such activity we can truly and simply say that it *is* the presence of the agent's will in the world.

We cannot *simply* say this of merely instrumental activities, valued merely extrinsically. To be sure, such activities can be said to reflect the agent's agency and practical identity in a thin sense.<sup>34</sup> After all, they are still intentional activities that proceed from the agent's instrumental reasoning in the pursuit of her own ends. However, this thin expression of agency is evidently insufficient for autonomy. It is present, for instance, in actions taken under coercion or duress—paradigm cases of unfreedom. The problem is that, insofar as it is *merely* instrumental, an activity expresses its agent's practical identity and will only in a qualified, privative way, for the full realization of her will in the world lies not within this action but outside it.

To illustrate the point, consider someone who performs routine labor for a living and is utterly indifferent to their job, doing it only to pay the bills. It would be a cruel joke to say of someone in such a position that their work expresses who they are as an agent and a valuer—that they choose their work

34 Thanks to Brookes Brown for pressing me on this point.

autonomously because what they want is precisely to perform such labor. Here is a striking passage from an interview with T, a call-center worker:

Would T be happy to think that his identity came from what he does all day? “I really hope not. I could not say enough how I hope not. I used to like who I was, and if this place is now my identity, then I don’t like myself. Literally, apart from the few people that I can sit and have a chat with and a gas with, the money is only just passable as the reason I come here. So, if the money changed, or certain people didn’t work here anymore, I can safely say I would probably be at the Job Centre looking.”<sup>35</sup>

T draws an immediate and natural connection between the fact that his work in some way fails to express his identity and the fact that it is a mere means to other ends. To be sure, he hedges a little, speaking only of what he “hopes” his identity is. But this ambivalence simply reflects the fact that there is *some* sense in which “who he is” is a matter of his actual doings—what he “does all day.” Evidently, the “self” constituted by these doings, though, is not one with which he identifies. And he explains why. The fundamental problem is not that he disavows any of his own first-order desires, or that these desires are in some other way “inauthentic.” It is rather that the relevant desires are merely extrinsic and so fail properly to express his practical identity. This is the sense in which he is alienated from them, and from the corresponding activities. We can thus do much of the philosophical work that we had hoped to do using higher-order desires, but while staying within the realm of first-order desire.<sup>36</sup>

We have arrived, then, at the Aristotelian-Marxian necessary condition of autonomy: one does something autonomously only if one does it for its own sake and not merely for the sake of further ends. This is not to say that all instrumental action is non-autonomous. There is nothing wrong as such with having to do one thing in order to do another. Indeed, many instrumental activities are loci of great and irreplaceable intrinsic value, and motivated in part by this value. The point is that an activity fails to be autonomous insofar as it is done *merely* instrumentally, for the sake of a further end alone.

I do not claim that doing something for its own sake is also a sufficient condition of autonomy. Our final ends are themselves vulnerable to distortion by the sorts of heteronomous psychological processes already familiar from the literature. The purpose of the new condition is to shed light on the comparatively

35 Biggs, *All Day Long*, 87.

36 In according a central role to the agent’s final ends, my account resonates with Gary Watson’s theory of autonomy, centered on a robust conception of the agent’s “values” (see Watson, “Free Agency”). Since Watson does not try to explain *external* unfreedom, however, the exact relation between my account and his is not straightforward.

poorly understood external dimension of autonomy. In this dimension, the question is not whether an agent has the right desires or whether they were formed in the right way, but whether her activity is an unqualified expression of her intrinsic desires.

### 3.2. *Making Sense of External Unfreedom*

The idea of autonomous activity as activity for its own sake promises an explanation of how coercion and duress normally compromise autonomy—an explanation that is structured, like our ordinary thought about these matters, by an opposition between a kind of desire and a kind of necessity.

Insofar as one is coerced to do something by means of a threat, one does it in order to avoid the threatened penalty. Volitionally coerced action is thus a species of instrumental action. Clearly, however, it must normally be a case of *merely* instrumental action. For, in general, if the coerced agent were already motivated to perform the action for its own sake, then it would be unnecessary to motivate her further by means of a threat. It follows that coerced action is normally non-autonomous. Its non-autonomy is a species of the non-autonomy of merely instrumental action in general, and can be explicated by saying that the coerced agent does not do what she really wants, in the sense that she does not do what she intrinsically wants, which is the robust kind of wanting that matters from the point of view of autonomy.

Much the same can be said of classical cases of duress, such as that of the sailor who throws goods overboard in a storm, the patient who chooses lifesaving surgery, or the person who abandons her house to flee a fire. Each of these seems to be a paradigm case of merely instrumental action: the agent does it intentionally but not for its own sake. In that sense, it is not something she really or truly wants to do.<sup>37</sup>

This explanation puts higher-order attitudes in their place. The agent acting merely instrumentally may well resent the way she is being motivated, whether by another person or by her circumstances. But we can now say, plausibly, that her resentment is a response to her unfreedom rather than the explanation of it. Thus, the Stoic slave does not become autonomous just by virtue of her equanimity. And even when an agent does resent being coerced, her non-autonomy is not a function of her resentment but its cause.

So much for the “desire” pole of the conceptual opposition. We also commonly say that agents acting under coercion or duress are forced, compelled,

37 Compare Aristotle’s discussion of the sailor’s action. Such actions, he says, seem voluntary inasmuch as they “are worthy of choice at the time when they are done,” but they cannot be called voluntary without qualification because “no one would choose any such act in itself” (Aristotle, *Nicomachean Ethics*, 1110a).

or necessitated to act as they do. Now we are in a position to explain what this means. Initially, I characterized my account in terms of a distinction between different sorts of desire, intrinsic and extrinsic. But to do something merely instrumentally, which is to act merely from an extrinsic desire, is to do it only because it is instrumentally necessary (other things equal) for the realization of one's further ends. It is to do what one *has* to do in order to do what one really wants to do. The distinctive kind of necessity that is the mark of external unfreedom, in other words, is *mere instrumental necessity*.

Unlike Frankfurt's attempt to account for the same thing, this explanation does not rely on the fiction of a psychologically irresistible desire, which would assimilate volitional coercion in general to special cases involving an overborne will. Moreover, the explanation locates the source of necessitation in the right place: not in the agent's own desire but in circumstances of her environment—those that account for the necessity of the means for the end. Finally, the relevant kinds of desire and necessity are not two separate, independently intelligible conditions but two sides of the same conceptual coin. To do something merely instrumentally is on the one hand not to do what one really wants to do (in the sense of what one intrinsically wants to do), and on the other hand—by the same token—to do merely what one is forced to do (in the sense of what one has to do for the sake of one's further ends).

The Aristotelian-Marxian account steers a path between subjectivism and objectivism. The subjectivist picture, guided by a preoccupation with psychological threats to autonomy, revolves around a contrast between authentic and inauthentic desires. Since this is essentially a division within the agent's mind, it is not surprising that the subjectivist picture struggles to make proper sense of external sources of unfreedom such as coercion and duress. The objectivist picture errs in the opposite direction. It takes more seriously the distinctiveness of external threats to autonomy, but it untethers this dimension of autonomy excessively from the agent's subjectivity. If duress is not understood fundamentally in conceptual opposition to some kind of desire, it proves difficult to understand how it can compromise autonomy.

The Aristotelian-Marxian picture does revolve around a conceptual opposition between (intrinsic) desire and (mere instrumental) necessity. This contrast remains tethered to the mind of the individual agent, in the form of her intrinsic desires or final ends. But at the same time, it situates the agent essentially within an external world, which is not just a place where we realize our desires but also a source of compulsions altogether alien to our desires. It is therefore no surprise that this alternative picture should be better placed than either purely subjectivist or purely objectivist approaches to make sense of external threats to our autonomy.

## 4. DOUBTS ABOUT NECESSITY

In the rest of the paper, I qualify and refine the basic Aristotelian-Marxian idea in important ways. I approach this task indirectly, by discussing some possible challenges to the view. In this section, I consider whether merely instrumental activity can really be said to be “necessitated” and hence unfree when there are other possible means to the same end or when that end is unimportant. I argue that mere instrumental necessity can be understood as a genuine form of necessity and a genuine source of unfreedom in both cases. I suggest, however, that the degree of this unfreedom is higher when the end in question is relatively important.

4.1. *Alternative Means*

Mere instrumental necessity is a species of instrumental necessity, or necessity for an end. It is rare, however, that a particular instrumental activity is the only possible means to its end; at least at a sufficiently specific level of description, there are usually bound to be alternatives. But if there are alternative means to a given end, it might seem that the means cannot be considered necessary for that end. Consider T, who works at the call center to pay his bills. If he could not pay his bills without working in some job or other, then we can say that he is forced *to work*: at this level of description, his activity is necessary for the further end (and done purely on account of this necessity). But if he could pay his bills by working at a fast-food restaurant instead, then it seems we cannot say that he is forced *to work at the call center*. He could have done something different and still achieved his end of paying the bills.

If this were right, then the core of my view could remain unchanged, but the relation between autonomy and necessity would be more complicated than I suggested in section 2.2. For I would have to say something like the following. An activity is autonomous only if it is done for its own sake, and not merely for the sake of further ends. But not all activities that fail to be autonomous are thereby *forced* or *necessitated*; or at any rate, they are not forced under all levels of description. A non-autonomous activity under a given description is forced only if there are no alternative means to the same end.

Perhaps we could put this in terms of a distinction between the absence of autonomy and the presence of unfreedom. Since T works purely for the sake of paying the bills, he does not work autonomously; and since he furthermore has no alternative means to this end, he is forced *to work* and does so unfreely. Similarly, T works at the call center purely for the sake of paying the bills, so he does not work at the call center autonomously; but since he has alternative means to the same end, at this level of description, he is not forced *to work at the*

call center and does not do so unfreely. His activity, at this level of description, falls short of autonomy but is not quite unfree.

I do not think any of this is obviously implausible. Yet there is something to be said for a more straightforward line of thought. First of all, it is important to see that there is a meaningful sense in which T's activity is genuinely necessary for its end even under the more specific description. While not necessary for the end without qualification, the activity is necessary for the end, *other things equal*, that is, assuming T does not take some other means to his end instead. This is a qualified but, I think, genuine sense of necessity, for it amounts to a genuine kind of counterfactual dependence.

By analogy, suppose a forest fire started because someone dropped a lit match on some dry leaves. The lit match may not have been necessary for the fire to start in the strictest possible sense. Perhaps, for example, the fire could equally have been started by the application of a flamethrower. Yet we are happy to say, I think, that the lit match being dropped was in some genuine sense a necessary condition of the fire. The reason we are happy to say this is that the latter event is counterfactually dependent on the former. That is, if the match had not been dropped, then the fire would not have started. This counterfactual clearly presupposes an other-things-equal condition, but it seems unproblematic.

The end of an instrumental action displays a similar counterfactual dependence on the action. That is, when someone does *X* for the sake of a further end *Y*, it is generally the case that she would not attain *Y* if she did not do *X*, other things equal. For instance, if T did not work at the call center, he would not get paid, other things equal. If the counterfactual dependence of the forest fire on the dropping of the match corresponds to a sense in which the latter was necessary for the former, the same seems true of the relation between an instrumental action and its further end. The action is necessary for its end, other things equal.<sup>38</sup>

Arguably, this qualified kind of necessity is the kind of necessity relevant for external unfreedom. A slave who is told to cook dinner might be left with a choice between cooking various specific meals. If she decides to cook stew, this action—cooking stew—is necessary for avoiding punishment only in the qualified, *ceteris paribus* sense: it is necessary for her end *if* she does not take some other means to her end—in particular, if she does not cook some other meal—instead. But to me, it seems absurd to deny that the slave is forced to cook the stew and cooks the stew unfreely. Accordingly, it seems to me on balance that the necessity in play when we say that someone is forced to do

38 Thanks to an anonymous reviewer for pressing me on this point.



something is not incompatible with the existence of alternative means to the same end. It is necessity, other things equal.<sup>39</sup>

#### 4.2. "Optional" Ends

Be that as it may, the end-relative nature of (mere) instrumental necessity raises a second worry. A low-wage call-center worker might work merely instrumentally in order to provide himself with food, shelter, and access to health care. A wealthy investment banker, with enough savings to retire comfortably, might also work merely instrumentally, perhaps in order to fund her dream of going on a commercial space flight or indulge some other "expensive taste." But there is surely a crucial difference here. Unlike the call-center worker, the banker seems to work for the sake of an end that is itself not essential but optional. Whereas food, shelter, and access to health care are necessities, the opportunity to go on a commercial space flight is a comparative luxury—an object, it is tempting to say, of *mere desire* as opposed to *need*. Is it not implausible, then, to say that the banker is forced to do her work? Surely a necessary condition of an optional end is itself optional, not necessary.

The idea that one of these ends but not the other is necessary without qualification, however, is a fiction that does not stand up to scrutiny. The call-center worker, for his part, could in principle choose to give up the end of obtaining food, shelter, and health care; it is not literally impossible for him to do this. To be sure, in giving up these ends, he would also be giving up the perceived goods he is after: survival itself, and everything to which survival is a means. But for the investment banker, too, to give up on the project of satisfying her expensive taste would be to give up the perceived good she is after. If there is a difference between the two cases, it does not consist in the fact that the call-center worker cannot do otherwise than pursue his end whereas the banker can do otherwise than pursue hers.

It does seem reasonable to say that the worker's end is likely to be more necessary for his good than the banker's end is for hers. Here, we are conceiving of the ends in question—survival and luxury space flight, respectively—also as means to some highly general end such as living a good life. This general good comes in degrees. And the point is that dying would likely result in a greater loss of it for the call-center worker than remaining permanently on Earth would for the banker. To the extent that this is true, it seems reasonable to say that the call-center worker's work is more necessary for his good, and that his work is in this sense necessitated more stringently. But given that he works *merely*

39 Note that this is consistent with allowing that the availability of (meaningfully different) alternative means is an important constituent of *global* autonomy.

instrumentally, this necessitation is a compulsion alien to his will. It therefore seems reasonable, too, to say that he works more unfreely than the banker does.

This picture has a subjectivist component, for I am suggesting that the degree to which a person's instrumental activity is necessitated is a matter of the degree to which her *perceived* good depends on it.<sup>40</sup> Conceivably if improbably, our banker might in fact place such a high value on luxury space travel that she would consider her life barely worth living without it. In that case, the work she did merely as a means to this further end would be necessitated and unfree to a very high degree. This strikes me as a plausible implication.

Notwithstanding this subjectivist component of my proposal, it is no accident that instrumental activities done for the sake of the things we usually call "basic necessities" are often necessitated to a distinctively high degree. This is not just because it is highly unusual for people to value luxuries such as space travel so obsessively that they would not consider their lives worth living without them. More importantly, it is because basic necessities tend to be highly *load bearing*: they tend to be necessary not just for this or that particular perceived good but for a very large range of perceived goods.<sup>41</sup>

For instance, food and shelter are necessary for survival, which is obviously necessary for very many of the things people might want to do with their lives—including, indeed, the enjoyment of luxuries such as flying to space for fun. Thus, even for the banker who greatly values this particular luxury, any work she does merely in order to secure her survival is likely to be still more necessary, and still more unfree, than the work she does merely in order to pay for her space flight. By the same token, we can say more generally that merely instrumental activity for the sake of procuring basic necessities tends to be highly unfree.

This claim rests on the following general principle: a merely instrumental activity is more stringently necessitated and hence more unfree, other things equal, when more is at stake for the agent. The principle also applies in an independently plausible way to classical cases of coercion and duress. Coercing someone by threatening her life makes her action more unfree than coercing her by threatening to destroy her stamp collection. There is greater unfreedom in leaving one's house to escape a fire than in fleeing to escape a rat infestation. I take it to be a strength of my account that it is able to vindicate and explain these comparative judgments of unfreedom.

Importantly, however, the difference in each case is altogether a matter of degree. For all that has been shown, there is no *categorical* difference between the call-center worker and the banker as far as the autonomy of their activity is

40 Thanks to Brookes Brown and to an anonymous reviewer for urging me to clarify this point.

41 Compare Rawls's notion of "primary goods" (*A Theory of Justice*, sec. 15).

concerned. In particular, it is not the case that the banker works autonomously. To be sure, for pragmatic reasons, we may shy away from rhetorically powerful words such as “forced” to describe cases like hers, where what is at stake for the agent is likely fairly trivial. Strictly speaking, though, both workers work unfreely, albeit to different degrees. Both work not under the guise of (intrinsic) desire but under the guise of (mere instrumental) necessity.

#### 5. TOO MUCH UNFREEDOM?

I have defended the Aristotelian-Marxian approach to freedom on grounds of general principle, as a straightforward interpretation of the idea that free activity is done from desire as opposed to necessity. I have also defended it on the basis of its explanatory power with respect to paradigm cases of unfreedom, for it makes sense of cases of duress and coercion that standard theories of autonomy struggle to illuminate. Yet the view considerably shrinks the domain of autonomous activity compared with competing theories, for if it is true, then merely instrumental activity is as such unfree. And this implication will strike many people as unacceptable. First, it might be argued, the implication is counterintuitive in itself, for many instrumental activities do not ordinarily strike us as unfree. Second, the implication may seem to devalue the concept of freedom, for many merely instrumental activities seem intuitively unobjectionable, so if they really are unfree, then their unfreedom cannot be of a kind worth caring about—or so one might argue.

Now, in general, that a view has counterintuitive implications is not necessarily a decisive objection against the view. Beyond paradigm cases, our intuitions are often unstable, subject to disagreement, and undoubtedly fallible. Sometimes, systematic theory should lead us to revise our intuitive judgments rather than vice versa. That said, I do not think that the Aristotelian-Marxian view requires as drastic a revision of our intuitions as one might think. In this section, I try to take some of the edge off, as it were, by arguing that the implications that appear most counterintuitive are largely confined to trivial cases where the relevant intuitions can be either accommodated or explained away.

##### 5.1. *Trivial Cases*

The intuitive case against the Aristotelian-Marxian picture is strongest, I think, with respect to very trivial sorts of action. At a sufficiently specific level of description, life seems full of merely instrumental activities.<sup>42</sup> In order to complete any number of tasks throughout my day, I have to push a variety of buttons—on

42 See, e.g., Jaeggi, *Alienation*, 207–8.

phones, elevators, washing machines, etc. Presumably, when I push such a button I normally do this not for its own sake but *merely* for the sake of its useful effect. On the Aristotelian-Marxian understanding, it follows that such activities are done unfreely—yet this implication seems strange in itself. Moreover, even if these activities are unfree, they seem obviously unobjectionable, which calls into question the normative significance of this category of supposed unfreedom.

I acknowledge that the unfreedom of such trivial activities seems untroubling in these sorts of cases. But this appearance has a straightforward explanation that is compatible with the Aristotelian-Marxian view. I have already argued that merely instrumental activities are less unfree if they serve unimportant *ends*; and presumably a lesser degree of unfreedom matters less than a greater degree, other things equal. I would now add that the unfreedom of merely instrumental activities likewise matters less, other things equal, when the *actions* themselves are insignificant, in the sense that they occupy a trivial place in the agent's life—as pushing the occasional button in the course of daily life surely does.

There are bound to be comparatively trivial instances of unfreedom on any plausible view of freedom and unfreedom. The distinction between those activities that really matter—those that play a significant role in our lives—and those that are comparatively unimportant is central to ethical and political thought in general. We presuppose it whenever we identify certain domains of human activity as salient subjects of freedom or autonomy in political philosophy: labor, sexuality, religion, political participation, and so on. As Charles Taylor puts it, “we make discriminations between obstacles as representing more or less serious infringements of freedom. And we do this, because we deploy the concept against a background understanding that certain goals and activities are more significant than others.”<sup>43</sup> The fact that a certain kind of unfreedom—such as the unfreedom of merely instrumental activity—is unimportant when the activities in question are trivial, then, hardly implies that it is not very important in other cases. The point is that good moral and political philosophy is partly a matter of focusing on the right things and asking the right questions.

The same considerations, I submit, can explain away our intuitive sense that these are not cases of unfreedom at all. The concept of unfreedom has its primary domain of application in the sphere of significant rather than trivial human activities, and in this primary domain it carries strong normative implications. When we apply the concept instead to very trivial activities where these implications are for all intents and purposes absent, we naturally feel reluctant to speak of genuine unfreedom, since some of the typical—normative—marks

43 Taylor contrasts the restrictions imposed by traffic lights with those imposed by a law limiting freedom of religion (“What’s Wrong with Negative Liberty,” 217–18).

of unfreedom are more or less missing. None of this shows, however, that the actions in question are without qualification *free*.

Certainly there are *specific* ways in which the activities in question can be said to be “free”—specific dimensions on which they may not be unfree. The ordinary, everyday sort of button pushing is free, for instance, in the specific sense that it is not coerced. This is the most important specific way in which an activity can be free or unfree, and in some contexts it will be all that matters. But uncoerced activities are not necessarily free *simpliciter*, or free in all ways. The way in which I have argued merely instrumental activities are unfree is the very generic sense that they are done under the guise not of desire properly speaking but of mere necessity, and therefore fail to be full and unqualified expressions of their agent’s will. Viewed in this light, the claim that even trivial sorts of merely instrumental activity are not free without qualification—that human freedom does not flourish in such activity—strikes me as an admittedly unfamiliar but ultimately principled and plausible upshot of our ordinary thought about freedom and compulsion.

### 5.2. *Nontrivial Cases*

What about less trivial cases of merely instrumental activity? It does not seem odd to me to speak of objectionable unfreedom in such cases. Consider what I take to be an importantly representative example, a different and rather more troubling case of button pushing. Here is Phil Stallings, describing his job as a spot welder in a car factory:

The welding gun’s got a square handle, with a button on the top for high voltage and a button on the bottom for low. . . . I stand in one spot, two- or three-foot area, all night. The only time a person stops is when the line stops. We do about thirty-two jobs per car, per unit. Forty-eight units an hour, eight hours a day. Thirty-two times forty-eight times eight. Figure it out. That’s how many times I push that button.<sup>44</sup>

The number is 12,288. Stallings’s button pushing would seem to be a nontrivial instance of merely instrumental activity if anything is. But the claim that it is unfree seems reasonable. Examples of this sort suggest, to my mind, that when a merely instrumental activity like button pushing comes to play a significant role in the agent’s life, its unfreedom is not unintuitive.<sup>45</sup>

44 Terkel, *Working*, 159.

45 For other arguments linking autonomy with meaningful work, see Schwartz, “Meaningful Work”; and Roessler, “Meaningful Work.”

Of course, there will be any number of cases between the extremes of my everyday button pushing and Stallings's all-day button pushing. Many ordinary household chores, for instance, probably fall somewhere in the middle. Here, too, I do not find the suggestion of unfreedom unintuitive; and in proportion as the activities are nontrivial, their unfreedom seems to me a genuine and obvious cause for concern. According to one study,

the proportion of working-class women spending more than 9 hours a week in washing and ironing dropped from 61 per cent before the introduction of the washing machine to 24 per cent afterwards.<sup>46</sup>

The work of washing and ironing, like factory labor, is arguably a form of toil: intrinsically unchoiceworthy labor. But if that is true, it does not seem a stretch to think of its reduction—other things equal—as a significant and highly desirable expansion of freedom. Of course, such work, like factory labor, is unequally distributed. This is an aggravating factor, and it partly explains why so many working-class women were spending absolutely so much time in this activity in the first place. But it does not explain the apparent unfreedom of the activity. On the contrary, the unequal distribution of toil is especially troubling precisely because it is a distribution of unfreedom.

Technology and the social reorganization of work can reduce the amount of extreme toil in human life. But perhaps merely instrumental activity, even of a nontrivial sort, cannot be altogether eliminated. Perhaps human life will always contain some measure of nontrivial labor done out of mere necessity. Both the Aristotle of book x of the *Ethics* and the Marx of volume 3 of *Capital* seem to have thought so. Aristotle admits that we cannot spend literally our whole lives in contemplation for its own sake, while Marx for his part suggests that the presence of a “realm of necessity” is an unavoidable fact “in all forms of society and under all possible modes of production.”<sup>47</sup> If this moderately pessimistic assessment of the human condition is right, then human life must always fall somewhat short of the full and unqualified realization of the ideal of freedom. All we can realistically aspire to is to approximate the ideal as closely as possible. But is there anything wrong with that aspiration? Perhaps Aristotle is on to something when he declares that

we must not follow those who advise us, being men, to think of human things, and, being mortal, of mortal things, but must, so far as we can, make ourselves immortal, and strain every nerve to live in accordance

46 Lynd and Lynd, *Middletown in Transition*, quoted in Gershuny, *Changing Times*, 67, quoted in Goodin et al., *Discretionary Time*, 75n29.

47 Marx, *Capital*, vol. 3, 959. Compare Veltman, *Meaningful Work*, ch. 5.

with the best thing in us; for even if it be small in bulk, much more does it in power and worth surpass everything.<sup>48</sup>

## 6. CONCLUSION

My aim in this paper has been to put on the table a neglected but attractive approach to understanding external freedom inspired by Aristotle and Marx. To do something autonomously, I have argued, one must do it for its own sake and not merely instrumentally. This necessary condition explains the sense in which doing something autonomously is a matter of doing what one wants. It also explains the distinctive and correlative sense in which externally unfree action is marked by alien necessitation.

As finite creatures, we will always live our lives partly in a realm of instrumental necessity, having to do one thing because we want to do another. There is nothing wrong with this condition *per se*: nothing necessarily wrong, for example, with having to cook in order to eat, or having to eat in order to live. Indeed, as I have mentioned, many instrumental activities are loci of great and irreplaceable intrinsic value. But the point is that our subjection to such necessity is compatible with autonomy only insofar as it amounts to more than *mere* necessity. To be autonomous, instrumental activity must be at the same time an object of intrinsic desire—an unqualified expression of our will in the world.

The Aristotelian-Marxian view is not without its difficulties. It has seemingly counterintuitive implications, since there are many merely instrumental activities that we may not usually think of as unfree. However, I have argued that the view is not as counterintuitive as it seems. In trivial cases, I suggested, the relevant intuitions can be either accommodated or explained away. In non-trivial cases, on the other hand, the view's radical implications seem plausible on reflection. Indeed, I take these implications to be a strength of the view. In addition to shedding light on traditional paradigm cases of coercion and duress, it fruitfully expands our philosophical understanding of autonomy by bringing into view distinctive species of external unfreedom.

Chief among these is the unfreedom of toil. Toil as such is unfree, not because it is necessarily coerced or done from especially dire economic need but because it is an impoverished, intrinsically worthless form of activity done only as a means to further ends. While the political implications of the Aristotelian-Marxian idea are beyond the scope of this paper, this thought might certainly give us pause. Perhaps we will want to reevaluate how well capitalist societies—which have brought us the steam engine and the computer but also

48 Aristotle, *Nicomachean Ethics*, 1177b.



factory labor and scientific management—live up to their promise of protecting and promoting the autonomy of the individual.<sup>49</sup>

Clemson University  
pbrixel@clemson.edu

## REFERENCES

- Arendt, Hannah. *The Human Condition*, 2nd ed. Chicago: University of Chicago Press, 1998.
- Aristotle. *Nicomachean Ethics*. In *The Complete Works of Aristotle*, edited by Jonathan Barnes, 1729–1867. Princeton, NJ: Princeton University Press, 1984.
- . *Politics*. In *The Complete Works of Aristotle*, edited by Jonathan Barnes, 1986–2129. Princeton, NJ: Princeton University Press, 1984.
- Biggs, Joanna. *All Day Long: A Portrait of Britain at Work*. London: Profile Books, 2015.
- Bratman, Michael E. “Identification, Decision, and Treating as a Reason.” *Philosophical Topics* 24, no. 2 (Fall 1996): 1–18.
- Christman, John. “Autonomy and Personal History.” *Canadian Journal of Philosophy* 21, no. 1 (March 1991): 1–24.
- . “Introduction.” In *The Inner Citadel: Essays on Individual Autonomy*, 3–23. Brattleboro, VT: Echo Point Books and Media, 2014.
- . “Liberalism and Individual Positive Freedom.” *Ethics* 101, no. 2 (January 1991): 343–59.
- . “Procedural Autonomy and Liberal Legitimacy.” In *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy*, edited by James Stacey Taylor, 277–98. Cambridge: Cambridge University Press, 2005.
- Cohen, G. A. “The Structure of Proletarian Unfreedom.” *Philosophy and Public Affairs* 12, no. 1 (Winter 1983): 3–33.
- Dworkin, Gerald. “Acting Freely.” *Noûs* 4, no. 4 (November 1970): 367–83.
- . “Autonomy and Behavior Control.” *The Hastings Center Report* 6, no. 1 (February 1976): 23–28.
- . “The Concept of Autonomy.” In *The Inner Citadel: Essays on Individual*

49 For extended and valuable feedback that improved this paper, I want to thank Amichai Amit, Anastasia Berg, Brookes Brown, Dan Brudney, Jacob Butcher, A. J. Julius, Ben Lawrence, Gabriel Lear, audiences at the Practical Philosophy Workshop and the Graduate Research Symposium at the University of Chicago, several anonymous reviewers, and especially Anton Ford and Claire Kirwin.

- Autonomy*, edited by John Christman, 54–62. Brattleboro, VT: Echo Point Books and Media, 2014.
- . *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press, 1988.
- Ekstrom, Laura Waddell. “A Coherence Theory of Autonomy.” *Philosophy and Phenomenological Research* 53, no. 3 (September 1993): 599–616.
- Epictetus. *The Discourses of Epictetus*. Translated by Elizabeth Carter and Robin Hard. London: Everyman, 1995.
- Frankfurt, Harry. “Coercion and Moral Responsibility.” In *Essays on Freedom of Action*, edited by Ted Honderich, 63–86. London: Routledge and Kegan Paul, 1973.
- . “The Faintest Passion.” *Proceedings and Addresses of the American Philosophical Association* 66, no. 3 (November 1992): 5–16.
- . “Freedom of the Will and the Concept of a Person.” *Journal of Philosophy* 68, no. 1 (January 1971): 5–20.
- . “Identification and Wholeheartedness.” In *The Importance of What We Care About*, 159–76. Cambridge: Cambridge University Press, 1988.
- Gershuny, Jonathan. *Changing Times: Work and Leisure in Post-Industrial Society*. Oxford: Oxford University Press, 2000.
- Goodin, Robert E., James Mahmud Rice, Antti Parpo, and Lina Eriksson. *Discretionary Time: A New Measure of Freedom*. Cambridge: Cambridge University Press, 1998.
- Hurka, Thomas. “Why Value Autonomy?” *Social Theory and Practice* 13, no. 3 (Fall 1987): 361–82.
- Jaeggi, Rahel. *Alienation*. Translated by Frederick Neuhouser and Alan E. Smith. Edited by Frederick Neuhouser. New York: Columbia University Press, 2014.
- Killmister, Suzy. *Taking the Measure of Autonomy: A Four-Dimensional Theory of Self-Governance*. Milton Park, UK: Routledge, 2018.
- Korsgaard, Christine M. “Two Distinctions in Goodness.” *Philosophical Review* 92, no. 2 (April 1983): 169–95.
- Kristinsson, Sigurdur. “The Limits of Neutrality: Toward a Weakly Substantive Account of Autonomy.” *Canadian Journal of Philosophy* 30, no. 2 (June 2000): 257–86.
- List, Christian, and Laura Valentini. “Freedom as Independence.” *Ethics* 126, no. 4 (July 2016): 1043–74.
- Lovett, Francis N. “Domination: A Preliminary Analysis.” *Monist* 84, no. 1 (January 2001): 98–112.
- Lynd, Robert S., and Helen Merrell Lynd. *Middletown in Transition: A Study in Cultural Conflict*. New York: Harcourt Brace, 1937.
- Marx, Karl. *Capital: A Critique of Political Economy*. Vol. 3. Translated by David

- Fernbach. London: Penguin, 1981.
- Mele, Alfred R. *Autonomous Agents: From Self-Control to Autonomy*. New York: Oxford University Press, 1995.
- Neely, Wright. "Freedom and Desire." *Philosophical Review* 83, no. 1 (January 1974): 32–54.
- Nozick, Robert. "Coercion." In *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, edited by Sidney Morgenbesser, Patrick Suppes, and Mary Terrell White. New York: St. Martin's Press, 1969.
- Oshana, Marina. "Personal Autonomy and Society." *Journal of Social Philosophy* 29, no. 1 (Spring 1998): 81–102.
- . *Personal Autonomy in Society*. Aldershot, UK: Ashgate Publishing, 2006.
- Pettit, Philip. *Republicanism: A Theory of Freedom and Government*. Oxford: Oxford University Press, 1997.
- Plamenatz, J. P. *Consent, Freedom and Political Obligation*. London: Oxford University Press, 1938.
- Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.
- Raz, Joseph. *The Morality of Freedom*. Oxford: Oxford University Press, 1986.
- Roessler, Beate. "Meaningful Work: Arguments from Autonomy." *Journal of Political Philosophy* 20, no. 1 (March 2012): 71–93.
- Schwartz, Adina. "Meaningful Work." *Ethics* 92, no. 4 (July 1982): 634–46.
- Tannenbaum, Julie. "Categorizing Goods." In *Oxford Studies in Metaethics*, vol. 5, edited by Russ Shafer-Landau, 257–94. Oxford: Oxford University Press, 2010.
- Taylor, Charles. "What's Wrong with Negative Liberty." In *Philosophical Papers*, 211–29. Cambridge: Cambridge University Press, 1985.
- Taylor, James Stacey. "Autonomy, Duress, and Coercion." *Social Philosophy and Policy* 20, no. 2 (July 2003): 127–55.
- . *Practical Autonomy and Bioethics*. New York: Routledge, 2009.
- Terkel, Studs. *Working: People Talk about What They Do All Day and How They Feel about What They Do*. New York: Random House, 1972.
- Thalberg, Irving. "Hierarchical Analyses of Unfree Action." *Canadian Journal of Philosophy* 8, no. 2 (June 1978): 211–26.
- Van Parijs, Philippe. *Real Freedom for All: What (If Anything) Can Justify Capitalism?* Oxford: Oxford University Press, 1995.
- Velleman, J. David. *The Possibility of Practical Reason*. Oxford: Oxford University Press, 2000.
- Veltman, Andrea. *Meaningful Work*. Oxford: Oxford University Press, 2016.
- Watson, Gary. "Free Agency." *Journal of Philosophy* 72, no. 8 (April 1975): 205–20.
- Zimmerman, David. "Making Do: Troubling Stoic Tendencies in an Otherwise Compelling Theory of Autonomy." *Canadian Journal of Philosophy* 30, no. 1 (March 2000): 25–54.